

TOPOLOGICAL ENTROPY OF RANDOM WALKS ON MAPPING CLASS GROUPS

HIDETOSHI MASAI

ABSTRACT. For any pseudo-Anosov diffeomorphism on a closed orientable surface S of genus greater than one, it is known by the work of Bers and Thurston that the topological entropy agrees with the translation distance on the Teichmüller space with respect to the Teichmüller metric. In this paper, we consider random walks on the mapping class group of S . The drift of a random walk is defined as the translation distance of the random walk. We define the topological entropy of a random walk and prove that it almost surely agrees with the drift on the Teichmüller space with respect to the Teichmüller metric.

1. INTRODUCTION

Let S be a closed orientable surface of genus ≥ 2 . According to the Nielsen-Thurston classification [Thu], every non-periodic irreducible automorphism of S is isotopic to a pseudo-Anosov diffeomorphism. Thurston proved that the topological entropy of any pseudo-Anosov diffeomorphism φ coincides with $\log \lambda_\varphi$ where λ_φ is the dilatation of φ (c.f. [FLP, Exposé 10]). Also by the work of Bers [Ber], $\log \lambda_\varphi$ is known to be equal to the translation distance of φ on the Teichmüller space with respect to the Teichmüller metric. The purpose of this paper is to demonstrate a “random version” of the work of Bers and Thurston. Let $\text{MCG}(S)$ denote the mapping class group of S . We consider the random walk on $\text{MCG}(S)$ which is determined by a probability measure μ on $\text{MCG}(S)$. This μ induces a probability measure \mathbb{P} on $\text{MCG}(S)^{\mathbb{Z}}$. Throughout the paper we assume that μ has finite first moment with respect to the Teichmüller metric and the support of μ generates a non-elementary subgroup of $\text{MCG}(S)$ (see Condition 2.2). Before stating the main theorem, we prepare several terminologies briefly. Formal definitions are given in §2. First, the topological entropy $h(\omega)$ of a sample path $\omega = (\omega_n) \in \text{MCG}(S)^{\mathbb{Z}}$ is defined using open coverings of S , similarly to the one for surface diffeomorphisms. This measures growth rate of the number of distinguishable orbits of the random walk. Next, Karlsson [Kar] proved that for \mathbb{P} -a.e $\omega = (\omega_n)$, the exponential growth rate of the length of the image $\omega_n(\alpha)$ of any simple closed curve α with respect to any metric always gives the same quantity, which is called the “Lyapunov exponent” $\lambda(\omega)$ of ω . Moreover, it is also proved that $\log \lambda(\omega)$ almost surely coincides with the drift $L_a(\omega)$ with respect to Thurston’s asymmetric Lipschitz metric on the Teichmüller space $\mathcal{T}(S)$ of S . Roughly speaking, the drift is the translation distance of ω . The goal of this paper is to show that those quantities are the same almost surely.

Theorem 1.1. *Let μ be a probability measure on $\text{MCG}(S)$ which satisfies Condition 2.2 and \mathbb{P} the probability measure on $\text{MCG}(S)^{\mathbb{Z}}$ induced by μ . For \mathbb{P} -a.e. $\omega \in$*

$\mathrm{MCG}(S)^{\mathbb{Z}}$, we have the following equality.

$$L_a(\omega) = \log \lambda(\omega) = h(\omega) = L_{\mathcal{T}}(\omega),$$

where $L_{\mathcal{T}}(\omega)$ is the drift with respect to the Teichmüller metric. These quantities are independent of ω and they are invariants of the random walk.

The strategy of the proof is similar to the one for pseudo-Anosov diffeomorphisms in [FLP, Exposé 9-10]. Indeed, $\log \lambda(\omega) \leq h(\omega)$ can be proved almost in the same way as the case of pseudo-Anosov diffeomorphisms. To prove the opposite inequality for pseudo-Anosovs, in [FLP], a subshift of finite type is associated to the dynamics of a pseudo-Anosov iteration by constructing so called a Markov partition. We will define a random subshift of finite type as a “random version” of a subshift of finite type (see §2.4). Then we will construct a semi-Markov partition of S which respects the dynamics of ω (see Definition 3.3) and associate to it a random subshift of finite type. The main difficulty, unlike pseudo-Anosov diffeomorphisms, is that for \mathbb{P} -a.e $\omega = (\omega_n) \in \mathrm{MCG}(S)^{\mathbb{Z}}$, some part of ω can be arbitrarily “bad”. For example, the orbit $\{\omega_n X\}$ of any point $X \in \mathcal{T}(S)$ may have large backtrack. On the other hand, for a pseudo-Anosov φ case, the fact that φ acts as a translation on a Teichmüller geodesic is implicitly used in [FLP]. To overcome the difficulty, in §3-4, we show that it suffices to observe only “good” elements in an orbit. The existence of such “good” elements follows from ergodic theorems.

To consider dynamics of ω on the surface, we need to take representatives of mapping classes. Let $\mathrm{Diff}^+(S)$ denote the space of orientation preserving diffeomorphisms on S . Let $w_n \in \mathrm{Diff}^+(S)$ be a representative of ω_n and $\mathbf{w} := (w_n)_{n \in \mathbb{Z}}$. Another difficulty occurs after taking representatives, that is, we can not use ergodic theorems. This is because we can not take representatives so that they are compatible with the shift maps on $\mathrm{MCG}(S)^{\mathbb{Z}}$, denoted θ . Notations (\mathbf{w}, n) are used instead of $\theta^n \omega$ to warn readers this issue. Our goal in §3 is to prove the following theorem.

Theorem 1.2. *There exist a random subshift of finite type $(\{\Sigma_A(\mathbf{w}, n)\}_{n \in \mathbb{Z}}, \sigma)$ such that the following diagram commutes for any $n \in \mathbb{Z}$.*

$$\begin{array}{ccc} \Sigma_A(\mathbf{w}, 0) & \xrightarrow{\sigma^n} & \Sigma_A(\mathbf{w}, n) \\ p(\mathbf{w}, 0) \downarrow & & \downarrow p(\mathbf{w}, n) \\ S & \xrightarrow{\omega_n^{-1}} & S \end{array}$$

where $p(\mathbf{w}, n) : \Sigma(\mathbf{w}, n) \rightarrow S$ is a continuous surjective map and σ is the shift map.

The topological entropy of σ in Theorem 1.2 can be defined as the growth rate of the number of cylinder sets of length n in $\Sigma_A(\mathbf{w}, 0)$. Let $h(\sigma)$ denote the topological entropy of σ . The fact $h(\sigma) \leq L_{\mathcal{T}}$ can be easily observed (Lemma 3.11). For pseudo-Anosovs, the facts of type Theorem 1.2 and $h(\sigma) \leq L_{\mathcal{T}}$ suffice to prove a theorem of type Theorem 1.1. However, we need to vary structures of S to construct a semi-Markov partition. Hence we need to discuss how structures, especially the Lebesgue number of a fixed open covering, vary as the steps. The Lebesgue numbers are discussed in §4.1, and the rest of §4 is devoted for a proof of $h(\omega) \leq h(\sigma)$.

2. PRELIMINARIES

In this section, we prepare terminologies and basic facts which we need to prove Theorem 1.1.

2.1. Teichmüller space. We briefly recall the Teichmüller spaces and related facts. Readers should refer [FM, FLP] for more details. Let S be a closed orientable surface of genus $g(S) > 1$. A *marked Riemann surface* is a pair of a Riemann surface \mathcal{X} and a homeomorphism, called a *marking*, $f : S \rightarrow \mathcal{X}$. Two marked Riemann surfaces $(\mathcal{X}_i, f_i : S \rightarrow \mathcal{X}_i)$, $(i = 1, 2)$ are said to be *Teichmüller equivalent* if there is a biholomorphic map $\phi : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ such that $\phi \circ f_1$ is homotopic to f_2 . The Teichmüller space $\mathcal{T}(S)$ of S is the space of marked Riemann surfaces modulo Teichmüller equivalence. Since each marking defines a complex structure on S by pullback, we often confuse a point $X \in \mathcal{T}(S)$ with a complex structure on S . The mapping class group $\text{MCG}(S)$ acts on $\mathcal{T}(S)$ so that for $\varphi \in \text{MCG}(S)$ with a representative $\psi \in \text{Diff}^+(S)$, $\varphi \cdot (\mathcal{X}, f : S \rightarrow \mathcal{X}) = (\mathcal{X}, f \circ \psi^{-1} : S \rightarrow \mathcal{X})$.

A holomorphic quadratic differential on $X \in \mathcal{T}(S)$ is a family of holomorphic maps $q = \{q_\alpha\}$ each defined on $z_\alpha(U_\alpha)$ of a complex chart $U_\alpha \subset X$, $z_\alpha : U_\alpha \rightarrow \mathbb{C}$ so that if $U_\alpha \cap U_\beta \neq \emptyset$ then

$$q_\beta(z_\beta) = q_\alpha \circ z_{\alpha\beta}(z_\beta) \cdot (z'_{\alpha\beta}(z_\beta))^2$$

where $z_{\alpha\beta} := z_\alpha \circ z_\beta^{-1}$. For $X \in \mathcal{T}(S)$, let $Q(X)$ denote the space of quadratic differentials. The vertical (resp. horizontal) trajectories of a quadratic differential q are curves $z(t)$ such that $q(z(t))z'(t)^2 \in \mathbb{R}_{>0}$ (resp. $\mathbb{R}_{<0}$). For each smooth arc τ , the transverse measures on the vertical and horizontal trajectories are defined by $\int_\tau |\Im q(z)^{1/2} dz|$ and $\int_\tau |\Re q(z)^{1/2} dz|$ respectively. Thus each $q \in Q(X)$ defines two measured foliations called *vertical* and *horizontal* foliations as the vertical and horizontal trajectories equipped with the transverse measures respectively. A theorem of Teichmüller says that given two points $X, Y \in \mathcal{T}(S)$, there exists a quasi-conformal map $T : X \rightarrow Y$ and quadratic differentials $q_X \in Q(X)$ and $q_Y \in Q(Y)$ such that the map T maps q_X to q_Y so that it stretches (resp. contracts) the horizontal (resp. vertical) foliations. The logarithm of the stretch factor coincides with the Teichmüller distance $d_{\mathcal{T}}(X, Y)$. By integrating the square root of a quadratic differential q , we have a singular Euclidean metric on X . With respect to the singular Euclidean metric, the length of a smooth arc τ' , denoted by $|\tau'|_q$, is equal to $\int_{\tau'} |q|^{1/2} dz$. For $q \in Q(X)$, we define the norm of q by $\|q\| := \int_X |q|$. Let $Q_1(X) := \{q \in Q(X) \mid \|q\| = 1\}$. We denote by $\mathcal{PMF}(S)$ the space of projective measured foliations. We consider the Thurston compactification $\bar{\mathcal{T}}(S) := \mathcal{T}(S) \cup \mathcal{PMF}(S)$ on which $\text{MCG}(S)$ acts continuously. By the work of Thurston, $\mathcal{PMF}(S)$ is homeomorphic to the sphere of dimension $6g(S) - 6$ [Thu]. We now recall the work of Hubbard-Masur.

Theorem 2.1 ([HM]). *The map $Q_1(X) \rightarrow \mathcal{PMF}(S)$ associating the equivalence class of the horizontal foliation to each $q \in Q_1(X)$ is a homeomorphism.*

Let $F, G \in \mathcal{PMF}(S)$ be transverse filling projective measured foliations. Let $\Gamma(F, G) \subset \mathcal{T}(S)$ denote the Teichmüller geodesic corresponding to a quadratic differential with horizontal and vertical foliation F and G respectively (see [GM] for the existence of such geodesics). A projective measured foliation is called *uniquely ergodic* if its supporting foliation admits only one transverse measure up to scale. Let $\mathcal{UE}(S) \subset \mathcal{PMF}(S)$ denote the space of uniquely ergodic foliations.

2.2. Random walk on group. Let G be a countable group and $\mu : G \rightarrow [0, 1]$ a probability measure. By \mathbb{Z}_+ (resp. \mathbb{Z}_-), we denote the space of positive (resp. negative) integers. For group elements $x_1, \dots, x_n \in G$, the subset

$$[x_1, \dots, x_n] := \{\omega = (\omega_i) \in G^{\mathbb{Z}_+} \mid \omega_i = x_i \text{ for } 1 \leq i \leq n\}$$

is called a *cylinder set*. The probability measure μ induces a probability measure \mathbb{P} on the space of sample paths $G^{\mathbb{Z}_+}$ so that

$$\mathbb{P}([x_1, \dots, x_n]) = \mu(x_0^{-1}x_1)\mu(x_1^{-1}x_2) \cdots \mu(x_{n-1}^{-1}x_n),$$

where x_0 is the initial element which is assumed to be the identity unless otherwise stated. We also consider the *reflected measure* $\check{\mu}(g) := \mu(g^{-1})$. Let $\check{\mathbb{P}}$ be the probability measure on $G^{\mathbb{Z}_-}$ induced by $\check{\mu}$. Then by the map $\omega = (\omega_n)_{n \in \mathbb{Z}} \mapsto ((\omega_n)_{n \in \mathbb{Z}_+}, (\omega_n)_{n \in \mathbb{Z}_-})$, the probability measure $\mathbb{P} \times \check{\mathbb{P}}$ induces a probability measure on $G^{\mathbb{Z}}$ which we again denote by \mathbb{P} . We define the *Bernoulli shift*, denoted by θ , as for any $k \in \mathbb{Z}$,

$$(\theta^k \omega)_n := \omega_k^{-1} \omega_{n+k}, \forall n \in \mathbb{Z}.$$

Recall that a subgroup of $\text{MCG}(S)$ is called *non-elementary* if it contains two pseudo-Anosov elements with disjoint fixed point sets in $\mathcal{PMF}(S)$. From now on, we consider the random walk on $\text{MCG}(S)$ which is determined by a probability measure μ which satisfies the following condition.

Condition 2.2. The probability measure $\mu : \text{MCG}(S) \rightarrow [0, 1]$ satisfies that

- μ has finite first moment with respect to the Teichmüller metric on $\mathcal{T}(S)$ i.e. for any $X \in \mathcal{T}(S)$, $\sum_{g \in \text{MCG}(S)} \mu(g) d_{\mathcal{T}}(X, gX) < \infty$, and
- the support of μ generates a non-elementary subgroup of $\text{MCG}(S)$.

2.3. Topological entropy, drift, and Lyapunov exponent. Let $\mathcal{A} = \{\mathcal{A}_i\}_{i \in I}$ and $\mathcal{B} = \{\mathcal{B}_j\}_{j \in J}$ be open coverings of S . Since S is compact, each open covering has a finite subcover. Let $N(\mathcal{A})$ denote the number of sets in a subcover of \mathcal{A} with minimal cardinality. \mathcal{B} is said to be a *refinement* of a cover \mathcal{A} , denoted $\mathcal{A} \prec \mathcal{B}$, if for any $B \in \mathcal{B}$, there is $A \in \mathcal{A}$ such that $B \subset A$. It can readily be seen that if $\mathcal{A} \prec \mathcal{B}$ then $N(\mathcal{A}) \leq N(\mathcal{B})$. We denote by $\mathcal{A} \vee \mathcal{B}$ the open cover $\{A_i \cap B_j\}_{i \in I, j \in J}$. For an open covering \mathcal{A} of S and a metric d on S , the *Lebesgue number* $\delta_d(\mathcal{A})$ with respect to d is defined to be

$$\inf_{x \in S} \sup\{r > 0 \mid B_r(x) \subset A \text{ for some } A \in \mathcal{A}\},$$

where $B_r(x)$ is the open ball centered at x of radius r with respect to d . Let $r > 0$ be less than the Lebesgue number of \mathcal{A} . Then the covering consisting of all open balls of radius r refines \mathcal{A} .

Definition 2.3 (Topological entropy. c.f.[AKM]). Let $\omega = (\omega_n)_{n \in \mathbb{Z}} \in \text{MCG}(S)^{\mathbb{Z}}$. We first choose an arbitrary representative $w_n \in \text{Diff}^+(S)$ of ω_n for each $n \in \mathbb{Z}$. Let $\mathbf{w} := (w_n)_{n \in \mathbb{Z}}$. For an open cover \mathcal{A} , let $N_n(\mathbf{w}, \mathcal{A}) := N(\mathcal{A} \vee w_1(\mathcal{A}) \vee \cdots \vee w_{n-1}(\mathcal{A}))$. We define

$$h(\mathbf{w}, \mathcal{A}) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log N_n(\mathbf{w}, \mathcal{A}).$$

Note that if $\mathcal{A} \prec \mathcal{B}$, then $h(\mathbf{w}, \mathcal{A}) \leq h(\mathbf{w}, \mathcal{B})$. The *topological entropy* of \mathbf{w} is

$$h(\mathbf{w}) := \sup_{\mathcal{A}} h(\mathbf{w}, \mathcal{A}),$$

where the supremum is taken over all open coverings of S . Finally we define

$$h(\omega) := \inf_{\mathbf{w}} h(\mathbf{w})$$

where the infimum is taken over all representatives of ω .

Remark 2.4. Unlike the definition of topological entropy of surface automorphisms, we do not take inverses. This is natural because when we consider random walks, we multiply new elements from the right.

We define the drift of random walks, which we may regard as a “translation distance” of the random walk.

Definition 2.5. Let (X, d_X) be a metric space on which $\text{MCG}(S)$ acts isometrically. Suppose the probability measure μ has finite first moment with respect to d_X , i.e.

$$\sum_{g \in \text{MCG}(S)} \mu(g) d_X(x, gx) < \infty,$$

where $x \in X$ is arbitrary. By Kingman’s subadditive ergodic theorem, the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} d_X(x, \omega_n x)$$

exists for \mathbb{P} -a.e. ω and this limit is independent of x and ω . This limit is called the *drift* of $\omega \in \text{MCG}(S)^{\mathbb{Z}}$ with respect to d_X .

Let d_a and $d_{\mathcal{T}}$ denote the distance on $\mathcal{T}(S)$ by Thurston’s Lipschitz metric and the Teichmüller metric respectively. We here recall the work of Choi-Rafi.

Theorem 2.6 ([CR, Theorem B]). *There is a constant c depending on the surface S and on δ such that for any X, Y in the δ -thick part of $\mathcal{T}(S)$, $d_{\mathcal{T}}(X, Y)$ and $d_a(X, Y)$ differ from one another by at most c .*

By Theorem 2.6, if the probability measure μ has finite first moment with respect to the Teichmüller metric, then it also has finite first moment with respect to Thurston’s Lipschitz metric. Let L_a (resp. $L_{\mathcal{T}}$) denote the drift of ω with respect to d_a (resp. $d_{\mathcal{T}}$). Since the drifts are also independent of the choice of base points, by taking a point in the thick part of $\mathcal{T}(S)$, we have $L_a = L_{\mathcal{T}}$ by Theorem 2.6. We let $L := L_a = L_{\mathcal{T}}$.

In [Kar], Karlsson proved the following.

Theorem 2.7 ([Kar]). *There exists λ such that for \mathbb{P} -a.e. $\omega \in \text{MCG}(S)^{\mathbb{Z}}$, for any isotopy class α of essential simple closed curves and Riemannian metric ρ of S ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log l_{\rho}(\omega_n^{-1} \alpha) = \log \lambda,$$

where $l_{\rho}(\alpha)$ denotes the infimum of the length of curves in α with respect to ρ . Moreover $\log \lambda$ coincides with L .

Note that for Theorem 2.7, we do not need to take a representative of ω . Following [DH], we call λ in Theorem 2.7 the *Lyapunov exponent* of the random walk.

Now we establish the following inequality.

Lemma 2.8. *Let λ be the Lyapunov exponent of the random walk determined by μ . For \mathbb{P} -a.e. ω , we have*

$$\log \lambda \leq h(\omega).$$

Proof. We first fix a hyperbolic metric ρ on S , a universal covering $\pi : \mathbb{H}^2 \rightarrow S$ and a representative $\mathbf{w} = (w_n)$ of ω . We also fix $p \in S$ and $\tilde{p} \in \pi^{-1}(p)$ in order to choose lifts \tilde{w}_n of w_n uniquely for all $n \in \mathbb{Z}$. In [FLP], a pseudo-Anosov diffeomorphism $\varphi : S \rightarrow S$ is discussed. One can prove the following lemma by exchanging φ^n with w_n^{-1} , and following the same argument as in [FLP].

Lemma 2.9 (c.f.[FLP, Lemma 10.8]). *For any $x, y \in \mathbb{H}^2$*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log d_\rho(\tilde{w}_n^{-1}x, \tilde{w}_n^{-1}y) \leq h(\mathbf{w}).$$

We may choose x and y in Lemma 2.9 to be the endpoints of a lift of geodesic representative of a simple closed curve α on S . Then since $l_\rho(w_n^{-1}\alpha) \leq d_\rho(\tilde{w}_n^{-1}x, \tilde{w}_n^{-1}y)$, we have $\log \lambda \leq h(\mathbf{w})$ for any representative \mathbf{w} of ω . \square

In order to prove $h(\omega) \leq L(= \log \lambda)$, we need a notion of random subshift of finite type.

2.4. Random subshift of finite type. We define a random subshift of finite type which we use to prove Theorem 1.1. Our goal is to associate a random subshift of finite type to a sample path $\omega \in \text{MCG}(S)^\mathbb{Z}$. Since we have to overcome certain difficulty which is described briefly in the introduction, we need to slightly modify the definition from the standard one (see e.g. [GK, Definition 3.9] for the standard one). The main difference is that we can only associate a random subshift of finite type to a representative \mathbf{w} of $\omega \in \text{MCG}^\mathbb{Z}(S)$. For later convenience, we use the notations with \mathbf{w} here.

Definition 2.10. Let $k(\mathbf{w}, \cdot) : \mathbb{Z} \rightarrow \mathbb{Z}$ be a function. Suppose we have a family of $k(\mathbf{w}, n) \times k(\mathbf{w}, n+1)$ matrices $A(\mathbf{w}, n)$ each of whose entry is 0 or 1. For any $n \in \mathbb{Z}$, let

$$\Sigma_k(\mathbf{w}, n) := \prod_{i \in \mathbb{Z}} \{1, 2, \dots, k(\mathbf{w}, i+n)\}.$$

We define the coordinate so that for each element $(x_i) \in \Sigma_k(\mathbf{w}, n)$, we have $x_i \in \{1, \dots, k(\mathbf{w}, i+n)\}$. A random subshift of finite type is a pair $(\{\Sigma_A(\mathbf{w}, n)\}_{n \in \mathbb{Z}}, \sigma)$ where

$$\Sigma_A(\mathbf{w}, n) := \{x = (x_i) \in \Sigma_k(\mathbf{w}) \mid A(\mathbf{w}, i+n)_{x_i, x_{i+1}} = 1 \text{ for all } i \in \mathbb{Z}\}$$

and $\sigma : \Sigma_A(\mathbf{w}, n) \rightarrow \Sigma_A(\mathbf{w}, n+1)$ is the standard left shift.

We consider the discrete topology on each $\{1, \dots, k(\mathbf{w}, n)\}$ and the product topology on $\Sigma(\mathbf{w}, n)$. Let $y_i \in k(\mathbf{w}, i+n)$ for $s \leq i \leq t$. The $(s, t)_n$ -cylinder set in $\Sigma_A(\mathbf{w}, n)$ of (y_i) is

$$\{x = (x_i) \in \Sigma_A(\mathbf{w}, n) \mid x_i = y_i, \text{ for all } s \leq i \leq t\}.$$

Let $\mathcal{C}_n(s, t)$ denote the family of $(s, t)_n$ -cylinder sets in $\Sigma_A(\mathbf{w}, n)$. Note that $\sigma(\mathcal{C}_n(s, t)) = \mathcal{C}_{n+1}(s-1, t-1)$.

3. CONSTRUCTION OF SEMI-MARKOV PARTITIONS

3.1. Birectangle partition. A semi-Markov partition with respect to $\omega \in \text{MCG}(S)^\mathbb{Z}$ is a sequence of partitions of the surface S by *birectangles* with certain condition so that it respects the dynamics of ω (see §3.2). We first construct a birectangle

decomposition, denoted $\mathcal{R}(F_+, F_-, \mathcal{X})$, from two transverse uniquely ergodic foliations $F_\pm \in \mathcal{UE}(S)$, and a marked Riemann surface $f : S \rightarrow \mathcal{X}$ which represents a point X on the Teichmüller geodesic $\Gamma(F_+, F_-)$. The Riemann surface structure \mathcal{X} lets us fix measured foliation representatives (\mathcal{F}_+, μ_+) and (\mathcal{F}_-, μ_-) of F_+ and F_- respectively so that (\mathcal{F}_+, μ_+) and (\mathcal{F}_-, μ_-) are the horizontal and vertical foliation of a holomorphic quadratic differential on \mathcal{X} of norm 1. Their preimages by f on S are also denoted by the same notations. Let $\text{Sing}(\mathcal{F})$ denote the set of singular points of \mathcal{F} . Note that with these representations, $\text{Sing}(\mathcal{F}_+) = \text{Sing}(\mathcal{F}_-)$.

Definition 3.1. A subset $R \subset S$ is called an $(\mathcal{F}_+, \mathcal{F}_-)$ -rectangle, or a *birectangle* if R is the image of some continuous map $\varphi : [0, 1] \times [0, 1] \rightarrow S$ such that

- $\varphi|_{(0,1) \times (0,1)}$ is an embedding, and
- for all $t \in [0, 1]$, $\varphi([0, 1] \times \{t\})$ (resp. $\varphi(\{t\} \times [0, 1])$) is a finite union of leaves and singularities of \mathcal{F}_+ (resp. \mathcal{F}_-), and in fact in one leaf if $t \in (0, 1)$.

We let $\text{int}(R) := \varphi((0, 1) \times (0, 1))$, $\partial_h R := \varphi([0, 1] \times \{0, 1\})$, $\partial_v R := \varphi(\{0, 1\} \times [0, 1])$, and $\partial R := \partial_h R \cup \partial_v R$.

A family of birectangles $\mathcal{R} = \{R_i\}$ is called a *birectangle partition* if

- (1) $\bigcup_i R_i = S$, and
- (2) $\text{int}(R_i) \cap \text{int}(R_j) = \emptyset$ for $i \neq j$.

For a singular measured foliation, we call a leaf which departs from a singularity a *singular leaf*. Any small neighborhood of a singular point is decomposed into several components by singular leaves. We call each component a *sector*. A *saddle connection* is a singular leaf which connects two singular points.

We will now construct a birectangle partition of S . We imitate the construction in [FLP, Exposé 9]. For each sector of \mathcal{F}_- of a singular point, we take a subarc of the singular leaf of \mathcal{F}_+ in the sector, which starts from the singular point and have μ_- measure 1. If \mathcal{F}_+ has a saddle connection and we can not take a singular leaf of μ_- measure 1, we instead take the whole saddle connection. Let $\tau' = \tau'(F_+, F_-, \mathcal{X}) \subset \mathcal{F}_+$ denote the family of such subarcs and saddle connections. Then for each singular leaf of \mathcal{F}_- , we take the shortest subarc that starts from a singular point and intersects every element of τ' which is not a saddle connection at least once. Similarly to before, we take whole saddle connections if there are no such subarcs. Let $\eta' = \eta'(F_+, F_-, \mathcal{X})$ denote the family of such subarcs and saddle connections. Then, for each $\alpha' \in \tau'$, we truncate the component of $\alpha' \setminus \eta'$ which contains $\partial\tau' \setminus \text{Sing}(\mathcal{F}_+)$ from α' , and denote by α the resulting arc. Note that saddle connections remain unchanged. Let $\tau = \tau(F_+, F_-, \mathcal{X}) := \{\alpha \mid \alpha' \in \tau'\}$. Then we extend each element of η' until it meets τ exactly once more. Let $\eta = \eta(F_+, F_-, \mathcal{X})$ denote the family of resulting subarcs. Then we let

$$\mathcal{R}(F_+, F_-, \mathcal{X}) := \{\overline{C} \mid C \text{ is a component of } S \setminus (\tau \cup \eta)\}.$$

Lemma 3.2. $\mathcal{R}(F_+, F_-, \mathcal{X})$ is a birectangle partition.

Proof. It suffices to prove that each element of $R \in \mathcal{R}(F_+, F_-, \mathcal{X})$ is a birectangle. If ∂R contains τ , then by construction R does not contain singular points in the interior. By the singular Euclidean structure determined by \mathcal{F}_\pm , we see that two components of $\partial R \cap \eta$ are parallel and in particular R is a birectangle. If there were R with $\partial R \cap \tau = \emptyset$, then ∂R must have contained a loop consisting of leaves of \mathcal{F}_- . However, since \mathcal{F}_- is uniquely ergodic, there are no such loops. \square

3.2. Semi-Markov partition. Let $\omega \in \text{MCG}(S)^{\mathbb{Z}}$ and $\mathbf{w} = (w_n)$ be a representative of ω . Our goal in this subsection is to construct a semi-Markov partition from a birectangle partition obtained in the previous subsection, so that it respects the dynamics of \mathbf{w} .

Definition 3.3. A sequence of birectangle partitions $\{\mathcal{R}_n\}_{n \in \mathbb{Z}}$ is a *semi-Markov partition* with respect to \mathbf{w} if for every $n \in \mathbb{Z}$,

- (M1) $w_n \partial_h \mathcal{R}_n \subset w_{n+1} \partial_h \mathcal{R}_{n+1}$, $w_n \partial_v \mathcal{R}_n \supset w_{n+1} \partial_v \mathcal{R}_{n+1}$, and
- (M2) for each $R_n \in \mathcal{R}_n$ and $R_{n+1} \in \mathcal{R}_{n+1}$, if $w_n R_n$ and $w_{n+1} R_{n+1}$ intersects, then the intersection is a single birectangle.

We call it a *semi-Markov partition* because to have $h(\omega) \leq L$, we further need estimates for the size of birectangles. We here carefully construct a semi-Markov partition so that it further satisfies certain estimates which we give in §4.

Recall that a Markov partition for a pseudo-Anosov diffeomorphism φ is constructed by using stable and unstable foliations F_s and F_u of φ . In terms of the Thurston compactification $\bar{\mathcal{T}}(S)$, these foliations are characterized as limits

$$\lim_{n \rightarrow \infty} \varphi^n X = F_s \text{ and } \lim_{n \rightarrow \infty} \varphi^{-n} X = F_u$$

where $X \in \mathcal{T}(S)$ is an arbitrary point. Kaimanovich-Masur proved that for the case of random walks, we have similar limits.

Theorem 3.4 ([KM, Theorem 2.2.4]). *Let μ be a probability measure which satisfies Condition 2.2. Then*

- (1) *There exists a unique μ -stationary probability measure ν on $\mathcal{PMF}(S)$ which is purely non-atomic and concentrated on $\mathcal{UE}(S)$.*
- (2) *For \mathbb{P} -a.e. $\omega \in \text{MCG}(S)^{\mathbb{Z}_+}$ and any $X \in \mathcal{T}(S)$, the sequence $\omega_n X$ converges in $\mathcal{PMF}(S)$ to a limit $F(\omega) \in \mathcal{UE}(S)$ and the distribution of the limits is given by the measure ν .*

We may apply Theorem 3.4 both to μ and $\check{\mu}$. We denote by $\check{\nu}$ the $\check{\mu}$ -stationary measure on $\mathcal{PMF}(S)$. For \mathbb{P} -a.e $\omega \in \text{MCG}(S)^{\mathbb{Z}}$, let

$$F_+(\omega) := \lim_{n \rightarrow +\infty} \omega_n X, \text{ and } F_-(\omega) := \lim_{n \rightarrow -\infty} \omega_n X.$$

Let $\Gamma(\omega)$ denote the Teichmüller geodesic $\Gamma(F_+(\omega), F_-(\omega))$. Note that by the definition of the Bernoulli shift θ , we have $F_+(\theta^n \omega) = \omega_n^{-1} F_+(\omega)$ and $F_-(\theta^n \omega) = \omega_n^{-1} F_-(\omega)$, and hence $\Gamma(\theta^n \omega) = \omega_n^{-1} \Gamma(\omega)$. We first fix X_0 on $\Gamma(\omega)$. By [KM, Lemma 1.4.3], the function $D : \mathcal{PMF}(S) \times \mathcal{PMF}(S) \rightarrow \mathbb{R}$, $(G_+, G_-) \mapsto d_{\mathcal{T}}(X_0, \Gamma(G_+, G_-))$ is continuous where it is defined. We fix open neighborhoods U_+ of $F_+(\omega)$ and U_- of $F_-(\omega)$ with the following condition.

Condition 3.5. The neighborhoods U_{\pm} satisfy

- U_+ and U_- have positive ν and $\check{\nu}$ measure respectively, and
- for any $G_+ \in U_+$ and $G_- \in U_-$, there is the Teichmüller geodesic $\Gamma[G_+, G_-]$.
- $D(U_+, U_-)$ is bounded from above by some constant $C > 0$.

The construction of semi-Markov partition in this section works for any U_{\pm} satisfying Condition 3.5. We give U_{\pm} which satisfy further condition that we need to prove Theorem 1.1 in §4.3. Finally let $\delta > 0$ be small enough so that the C -neighborhood of X_0 is contained in the δ -thick part of $\mathcal{T}(S)$.

We now choose points in $\mathcal{T}(S)$ to construct a semi-Markov partition. See Figure 1 for a schematic picture. First, we choose X'_n to be a closest point to X_0 on $\Gamma(\theta^n\omega)$. For n positive, we define $X''_n \in \Gamma(\theta^n\omega)$ inductively by

$$(1a) \quad X''_n := \begin{cases} X'_n & \text{if } F_+(\theta^n\omega) \in U_+ \text{ and } F_-(\theta^n\omega) \in U_- \\ \omega_n^{-1}\omega_{n-1}X''_{n-1}, & \text{otherwise.} \end{cases}$$

We define X''_n for negative n similarly. We then define $\varepsilon : \mathbb{Z} \rightarrow \{0, 1\}$ as follows. For positive n , we set $\varepsilon(n) = 1$ if $d(X_0, \omega_n X''_n) > d(X_0, \omega_i X''_i)$ for all $0 \leq i < n$, and $\varepsilon(n) = 0$ for otherwise. For negative n , ε is defined similarly. We set $\varepsilon(0) := 1$. Then for n positive, we define X_n inductively

$$X_n := \begin{cases} X''_n & \text{if } \varepsilon(n) = 1 \\ \omega_n^{-1}\omega_{n-1}X_{n-1}, & \text{if } \varepsilon(n) = 0. \end{cases}$$

We define X_n for negative n similarly. These X_n are in the δ -thick part and $\omega_n X_n$ are located according to the order of n on $\Gamma(\omega)$. Even with this modification, the distance between $\omega_n X_0$ and $\omega_n X_n$ grows sublinearly.

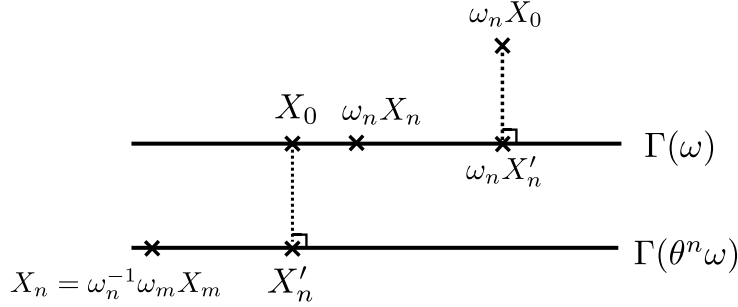


FIGURE 1. Positions of X_n . The symbol of right angle is used to mean a closest point projection. Where $m := \max\{k \in \mathbb{Z}_+ \mid k < n \text{ and } \varepsilon(m) = 1\}$.

Lemma 3.6. *For above ω and $\{X_n\}$,*

$$\lim_{n \rightarrow \pm\infty} \frac{1}{n} d_{\mathcal{T}}(\omega_n X_0, \omega_n X_n) = 0.$$

Proof. Since one can prove the statement for negative n similarly, we assume that n is positive. Then since ν and $\tilde{\nu}$ are independent,

$$\mathbb{P}(\eta = (\eta_m) \in \text{MCG}(S)^{\mathbb{Z}} \mid F_+(\eta) \in U_+ \text{ and } F_-(\eta) \in U_-) = \nu(U_+) \tilde{\nu}(U_-) > 0.$$

Hence by the ergodic theorem

$$\{n \in \mathbb{Z} : F_+(\theta^n\omega) \in U_+ \text{ and } F_-(\theta^n\omega) \in U_-\}$$

has positive density. We now recall the work of Tiozzo.

Theorem 3.7 ([Tio, Theorem 18]). *For \mathbb{P} -a.e. $\omega \in \text{MCG}(S)^{\mathbb{Z}}$, let Γ denote the Teichmüller geodesic $\Gamma(\omega)$ with parametrization by arc length and $\Gamma(0) = X_0$. Then we have*

$$\lim_{n \rightarrow \pm\infty} \frac{1}{n} d_{\mathcal{T}}(\omega_n X_0, \Gamma(Ln)) = 0$$

where L is the drift with respect to the Teichmüller metric.

Let

$$l_n(\omega) := \max\{k \mid \text{for all } 0 \leq n-k \leq i \leq n, F_+(\theta^i \omega) \notin U_+ \text{ or } F_-(\theta^i \omega) \notin U_-\}.$$

Since $l_n(\omega) \leq n$, l_n is an integrable function. Hence Kingman's subadditive ergodic theorem implies for \mathbb{P} -a.e. ω ,

$$\lim_{n \rightarrow \infty} l_n(\omega)/n = 0.$$

Then, we first estimate $d_{\mathcal{T}}(\Gamma(Ln), \omega_n X_n'')$. Let $\epsilon > 0$. By above observations, we may suppose that for large enough n , we have $d_{\mathcal{T}}(\Gamma(Lm), \omega_m X_m') \leq m\epsilon \leq n\epsilon$ where $m := n - l_n(\omega)$, and $l_n(\omega) \leq n\epsilon$. Then by definition, we have

$$\begin{aligned} d_{\mathcal{T}}(\Gamma(Ln), \omega_n X_n'') &\leq d_{\mathcal{T}}(\Gamma(Ln), \Gamma(Lm)) + d_{\mathcal{T}}(\Gamma(Lm), \omega_m X_m') \\ &\leq (L+1)n\epsilon \end{aligned}$$

Hence we have

$$\lim_{n \rightarrow \infty} d_{\mathcal{T}}(\Gamma(Ln), \omega_n X_n'')/n = 0.$$

Now we estimate $d_{\mathcal{T}}(\omega_n X_n'', \omega_n X_n)$. If $X_n'' \neq X_n$, then there exists m such that $d_{\mathcal{T}}(X_0, \omega_m X_m'') > d_{\mathcal{T}}(X_0, \omega_n X_n'')$. Suppose m is chosen to be maximum with this property so that $\omega_m X_m'' = \omega_n X_n$. Let $\epsilon > 0$. We may suppose that n is large enough so that $d_{\mathcal{T}}(\Gamma(Lk), \omega_k X_k'') \leq k\epsilon$ for $k \in \{n, m\}$. Then we have two cases; $d_{\mathcal{T}}(X_0, \Gamma(Ln)) > d_{\mathcal{T}}(X_0, \omega_m X_m'')$ or $d_{\mathcal{T}}(X_0, \Gamma(Ln)) \leq d_{\mathcal{T}}(X_0, \omega_m X_m'')$. Since $Lm < Ln$, in both cases we have,

$$d_{\mathcal{T}}(\omega_n X_n'', \omega_n X_n) = d_{\mathcal{T}}(\omega_n X_n'', \omega_m X_m'') \leq (n+m)\epsilon \leq 2n\epsilon.$$

Hence we have

$$\lim_{n \rightarrow \infty} d_{\mathcal{T}}(\omega_n X_n'', \omega_n X_n)/n = 0.$$

By Theorem 3.7 and the triangle inequality, we have the conclusion. \square

Let $f : S \rightarrow \mathcal{X}_0$ be a representative of X_0 . Since each $\omega_n X_n$ is on $\Gamma(\omega)$, there is the Teichmüller map T_n that stretches $F_+(\omega)$ and contracts $F_-(\omega)$ such that $f : S \rightarrow \mathcal{X}_n := T_n(\mathcal{X}_0)$ represents $\omega_n X_n$. Then let $\mathcal{R}'_n := \mathcal{R}(F_+(\omega), F_-(\omega), \mathcal{X}_n)$, $\tau_n := \tau(F_+(\omega), F_-(\omega), \mathcal{X}_n)$ and $\eta_n := \eta(F_+(\omega), F_-(\omega), \mathcal{X}_n)$. We denote the corresponding measured foliation representatives of $F_+(\omega)$ and $F_-(\omega)$ by $(\mathcal{F}_+(\omega, n), \mu_+(\omega, n))$ and $(\mathcal{F}_-(\omega, n), \mu_-(\omega, n))$ respectively. Since $\omega_n X_n$ are on $\Gamma(\omega)$, by the definition of ε , $\{w_n^{-1} \mathcal{R}'_n\}$ satisfies (M1). We need to decompose each birectangles in $w_n^{-1} \mathcal{R}'_n$ further to have a partition which satisfies (M2).

Given two birectangle partitions $\mathcal{R}, \mathcal{R}'$ with $\partial_h \mathcal{R} \subset \partial_h \mathcal{R}'$ and $\partial_v \mathcal{R} \supset \partial_v \mathcal{R}'$, let $\mathcal{R} \vee \mathcal{R}'$ denote the birectangle partition we get by cutting S by $\partial_h \mathcal{R}' \cup \partial_v \mathcal{R}$. Let $0 < i < j < k$ be indices which satisfy

- (1) $\varepsilon(i) = \varepsilon(j) = \varepsilon(k) = 1$, and
- (2) $\varepsilon(l) = 0$ for all $i < l < j$ or $j < l < k$,

We define $\mathcal{R}_j := w_j^{-1}(\mathcal{R}'_i \vee \mathcal{R}'_j \vee \mathcal{R}'_k)$. We note that $\mathcal{R}'_i \vee \mathcal{R}'_j \vee \mathcal{R}'_k$ is equal to

$$\{\bar{C} \mid C \text{ is a component of } S \setminus (w_k \tau_k \cup w_i \eta_i)\}.$$

For $n > 0$ with $\varepsilon(n) = 0$, let m be the largest integer which is less than n and $\varepsilon(m) = 1$. We define $\mathcal{R}_n := w_n^{-1} w_m \mathcal{R}_m$. For negative n , \mathcal{R}_n is defined similarly. By the construction, $\{\mathcal{R}_n\}$ still satisfies (M1).

Lemma 3.8. $\{\mathcal{R}_n\}$ is a semi-Markov partition with respect to $\mathbf{w} = (w_n)$.

Proof. If $\varepsilon(n) = 0$, the condition (M2) is apparently satisfied. Hence it suffices to prove for m and n with

- $\varepsilon(m) = \varepsilon(n) = 1$ and
- $\varepsilon(l) = 0$ for all $m < l < n$,

that for each $R_m \in \mathcal{R}_m$ and $R_n \in \mathcal{R}_n$, the intersection $w_m R_m \cap w_n R_n$ is either empty or a single birectangle. Since $\{\mathcal{R}_n\}$ satisfies (M1), we see that if the intersection $w_m R_m \cap w_n R_n \neq \emptyset$, it is a family of birectangles. Note that each birectangle in $w_m \mathcal{R}_m$ or $w_n \mathcal{R}_n$ is a subset of a component of $R'_m \cap R'_n$ for some $R'_m \in \mathcal{R}'_m$ and $R'_n \in \mathcal{R}'_n$. From each component R' of the intersection $R'_m \cap R'_n$, a birectangle $w_m R_m \in w_m \mathcal{R}_m$ (resp. $w_n R_n \in \mathcal{R}_n$) is obtained by decomposing R' vertically by leaves of F_- (resp. horizontally by leaves of F_+). Hence each $w_m R_m \cap w_n R_n$ is connected. Thus (M2) follows. \square

3.3. Symbolic dynamics. We now associate a random subshift of finite type to the representative \mathbf{w} of ω by using the semi-Markov partition $\{\mathcal{R}_n\}$ constructed in §3.2. Let $k(n, \mathbf{w})$ denote the number of birectangles in \mathcal{R}_n . We label birectangles in \mathcal{R}_n by $R_1^n, R_2^n, \dots, R_{k(n, \mathbf{w})}^n$. We define $k(n, \mathbf{w}) \times k(n+1, \mathbf{w})$ matrices $A(\mathbf{w}, n) = (a_{i,j}^n)$ by setting $a_{i,j}^n = 1$ if $w_n(\text{int}(R_i^n)) \cap w_{n+1}(\text{int}(R_j^{n+1})) \neq \emptyset$ and $a_{i,j}^n = 0$ for otherwise. Let $(\{\Sigma_A(\mathbf{w}, n)\}_{n \in \mathbb{Z}}, \sigma)$ be the random subshift of finite type with respect to $\{A(\mathbf{w}, n)\}_{n \in \mathbb{Z}}$. Then each element in $\Sigma_A(\mathbf{w}, n)$ corresponds to a point in S .

Lemma 3.9. *For any n and $\mathbf{b} = (b_i) \in \Sigma_A(\mathbf{w}, n)$,*

$$\bigcap_{i=-\infty}^{\infty} w_{i+n}(\text{int}(R_{b_i}^{i+n}))$$

determines a single point in S .

Proof. Let us fix $\mathbf{b} = (b_i) \in \Sigma_A(\mathbf{w}, n)$. By the properties (M1) and (M2) of semi-Markov partitions, we have that for each m , $C_m := \bigcap_{i=-m}^m w_{i+n}(\text{int}(R_{b_i}^{i+n}))$ is a birectangle with exactly one component. We consider the singular Euclidean metric that determines the point $\omega_n X_n$ on $\Gamma(\omega)$. Let Γ denote $\Gamma(\omega)$ with parametrization by arc length so that $\Gamma(0) = \omega_n X_n$. We will prove that the diameter of C_m converges to 0 as $m \rightarrow -\infty$. By Lemma 3.6, we see that points $\omega_m X_m$ for negative m are close to $\Gamma(L(n-m))$. To construct $\{\mathcal{R}_m\}$, we considered arcs on $F_+(\omega, m)$ of $\mu_+(\omega, m)$ measure 1 which is $\mu_+(\omega, n)$ measure almost equal to $1/\exp(L(n-m))$ by Lemma 3.6 and Theorem 3.7. Hence the horizontal diameter of C_m converges to 0 as $m \rightarrow -\infty$. On the other hand, for m positive the arcs $\tau(F_+(\omega), F_-(\omega), \mathcal{X}_m)$ travel on singular leaves of F_+ longer as m increases. Since each infinite singular leaf is dense, it follows that the vertical diameter converges to 0 as $m \rightarrow \infty$. Thus we have a point on \mathcal{X}_n . Finally by the marking $f \circ w_n : S \rightarrow \mathcal{X}_n$, we fixed above, we have a point on S . \square

By Lemma 3.9, we define $p(\mathbf{w}, n) : \Sigma_A(\mathbf{w}, n) \rightarrow S$.

Lemma 3.10 (c.f. [FLP, §10.4]). *The map $p(\mathbf{w}, n)$ is continuous and surjective.*

Proof. Since the image of a long cylinder set of $\Sigma_A(\mathbf{w}, n)$ is contained in a small birectangle, $p(\mathbf{w}, n)$ is continuous. Let $V_j = \bigcup_{i=1}^{k(j, \mathbf{w})} w_j(\text{int}(R_i^j))$. For each j , V_j is an open dense set. Then by the Baire category theorem, $U := \bigcap_{j \in \mathbb{Z}} V_j$

is dense. Each $x \in U$ is contained in $w_{j+n}(\text{int}(R_{b_j}^{j+n}))$ for some b_j for every $j \in \mathbb{Z}$. Let $\mathbf{b} = \{b_j\}_{j \in \mathbb{Z}} \in \Sigma_A(\mathbf{w}, n)$. We have $p(\mathbf{w}, n)(\mathbf{b}) = x$, which implies $U \subset p(\mathbf{w}, n)(\Sigma_A(\mathbf{w}, n))$. Since U is dense and $\Sigma_A(\mathbf{w}, n)$ is compact, $p(\mathbf{w}, n)$ is surjective. \square

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. Note that $T_n : \mathcal{X}_0 \rightarrow \mathcal{X}_n$ changes only the metric and does not change the image. Since $p(\mathbf{w}, n) : \Sigma_A(\mathbf{w}, n) \rightarrow S$ is defined by using $f \circ w_n : S \rightarrow \mathcal{X}_n$, we have $p(\mathbf{w}, n) \circ \sigma^n = w_n^{-1} \circ p(\mathbf{w}, 0)$. \square

We now consider the topological entropy of the shift map $\sigma : \Sigma_A(\mathbf{w}, n) \rightarrow \Sigma_A(\mathbf{w}, n+1)$. In order to prove Theorem 1.1, it suffices to prove that growth ratio of the number of elements of cylinder sets.

Lemma 3.11. *Let $(\{\Sigma_A(\mathbf{w}, n)\}_{n \in \mathbb{Z}}, \sigma)$ be the random subshift of finite type defined above. Then for any $K \in \mathbb{Z}$,*

$$\limsup_{m \rightarrow \infty} \frac{\log N(\mathcal{C}_n(K, K+m))}{m} \leq L,$$

where L is the drift of ω with respect to the Teichmüller metric.

Proof. Note that by the property (M1) and (M2) of semi-Markov partitions, the intersections $\mathcal{R}_K \vee \cdots \vee \mathcal{R}_{K+m}$ is also a birectangle partition. For a given birectangle partition \mathcal{R} , let $N(\mathcal{R})$ denotes the number of birectangles. By the map $p(\mathbf{w}, n)$, we see that

$$N(\mathcal{C}_n(K, K+m)) = N(\mathcal{R}_K \vee \cdots \vee \mathcal{R}_{K+m}).$$

Hence we will give a bound of $N(\mathcal{R}_K \vee \cdots \vee \mathcal{R}_{K+m})$. Let c_K be the shortest horizontal length of birectangles in \mathcal{R}_K measured by $\mu_-(\omega, n)$. Let L_m denote the maximum of $\mu_-(\omega, n)$ measures of the arcs $\tau_{K+m} := \tau(F_+(\omega), F_-(\omega), \mathcal{X}_{K+m})$. Each arc in τ_{K+m} cuts birectangles in \mathcal{R}_K at most L_m/c_K times. The number of singular leaves of $F_+(\mathbf{w}, n)$ is bounded from above by some constant D which depends only on S . Hence $N(\mathcal{R}_K \vee \cdots \vee \mathcal{R}_{K+m})$ is at most $D \cdot L_m/c_K$. By Lemma 3.6 and Theorem 3.7,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \log L_m = L,$$

which implies

$$\limsup_{m \rightarrow \infty} \log(N(\mathcal{R}_K \vee \cdots \vee \mathcal{R}_{K+m})) \leq L.$$

\square

4. PROOF OF THE MAIN THEOREM

In §3, we have constructed a semi-Markov partition $\{\mathcal{R}_n\}$ for any representative of \mathbb{P} -a.e. $\omega \in \text{MCG}(S)^{\mathbb{Z}}$. In this section, we will prove that for \mathbb{P} -a.e. $\omega \in \text{MCG}(S)^{\mathbb{Z}}$, we can find a representative $\mathbf{w} = (w_i)$ of ω such that $h(\mathbf{w}) \leq L$. In the case of pseudo-Anosov diffeomorphisms, facts of type Lemma 3.10 and 3.11 suffice to prove that the topological entropy and the translation distance on the Teichmüller space agree. This is because to construct a Markov partition for a pseudo-Anosov diffeomorphism, we only need to use a single point in $\mathcal{T}(S)$. On the other hand, for random walks, we need to use different X_n 's for each n . One of the main difficulty caused for above reason is that the Lebesgue number of a given open covering \mathcal{A} varies depending on the metric. In §4.1, we first give a suitable asymptotic bound

for the Lebesgue number. Every argument so far works for any U_{\pm} satisfying 3.5. In §4.2-4.3, the neighborhoods U_{\pm} of $\mathcal{F}_{\pm}(\omega)$ which we need to prove Theorem 1.1 are given.

4.1. Bound for the Lebesgue number. To have a bound of the Lebesgue number, we first observe how singular Euclid structures may change in the δ -thick part of $\mathcal{T}(S)$.

Lemma 4.1. *There exists $B = B(S, \delta)$ such that the following holds. Let X be in the δ -thick part of $\mathcal{T}(S)$, and $q_1, q_2 \in Q_1(X)$. Then the singular Euclidean metric associated to q_1 and q_2 are B -bi-Lipschitz.*

Proof. As pointed out in [FLP, Lemma 9.22], any two singular Euclidean metrics are bi-Lipschitz with some bi-Lipschitz constant. Since by Theorem 2.1, $Q_1(X)$ is homeomorphic to $\mathcal{PMF}(S)$, a compact space, we see that for any $X \in \mathcal{T}(S)$, there exists $B(X) > 0$ such that any two singular Euclidean metrics corresponding to elements in $Q_1(X)$ are $B(X)$ -bi-Lipschitz. Since $B(X)$ varies continuously and the δ -thick part of the moduli space of S is compact, we have a desired bound. \square

Recall that the semi-Markov partition $\{\mathcal{R}_n\}$ is defined on $X_n \in \mathcal{T}(S)$ with representative $f \circ w_n : S \rightarrow \mathcal{X}_n$. Then there is a quadratic differential $q_n \in Q(\mathcal{X}_0)$ that is the initial quadratic differential of the Teichmüller geodesic connecting X_0 and X_n . Let \mathcal{X}'_n denote the complex structure we get by stretching (resp. contracting) horizontal (resp. vertical) foliation of q_n so that it gives the same point as X_n in $\mathcal{T}(S)$. Let T'_n be the corresponding Teichmüller map. Since two markings $f \circ w_n : S \rightarrow \mathcal{X}_n$ and $f : S \rightarrow \mathcal{X}'_n$ gives the same point in the Teichmüller space, there is a biholomorphic map $\phi : \mathcal{X}'_n \rightarrow \mathcal{X}_n$ so that $\phi \circ f \circ w_n$ is homotopic to f . Hence by homotopy, we may suppose $w_n = f^{-1} \circ \phi^{-1} \circ f$. From now on, we use these representations and let $\mathbf{w} = (w_n)$. We now fix an open covering \mathcal{A} of S . Let q'_n be the quadratic differential on S determined by \mathcal{X}'_n and $\Gamma(\theta^n \omega)$. For a quadratic differential q , we denote by $\delta(q)$ the Lebesgue number of \mathcal{A} with respect to the singular Euclidean metric defined by q . By the choice of the representative $\mathbf{w} = (w_n)$, $\delta(q'_n)$ is equal to the Lebesgue number of $w_n \mathcal{A}$ with respect to the quadratic differential determined by $\Gamma(\omega)$ and \mathcal{X}_n that we used to construct $\mathcal{R}(F_+(\omega), F_-(\omega), \mathcal{X}_n)$.

Lemma 4.2. *For \mathbb{P} -a.e. ω , the $\{q'_n\}$ defined above satisfies*

$$\lim_{n \rightarrow \infty} \frac{-\log \delta(q'_n)}{n} = 0.$$

Proof. We first note that if two singular Euclidean metrics determined by q and q' are B -bi-Lipschitz, then we have $\delta(q)/\delta(q') \leq B$. Since we have chosen X_n so that they are in the δ -thick part, we have $\delta(q'_0)/\delta(q_n) \leq B$ and $\delta(T'_n(q_n))/\delta(q'_n) \leq B$ by Lemma 4.1. By the definition of T'_n , the ratio $\delta(q_n)/\delta(T'_n(q_n))$ is bounded from above by $\exp(d_{\mathcal{T}}(X_0, X_n))$. Hence

$$1/\delta(q'_n) \leq B^2 \exp(d_{\mathcal{T}}(X_0, X_n))/\delta(q'_0).$$

Therefore, by Lemma 3.6, we have the conclusion. \square

4.2. ν -measures of neighborhoods of $F_+(\omega)$. The goal of this subsection is the following proposition. The measure ν is from Theorem 3.4.

Proposition 4.3. *For \mathbb{P} -a.e. ω , any open neighborhood U of $F_+(\omega)$ has positive ν -measure.*

We first recall the curve graphs and shadows. The *curve graph* of S , denoted $\mathcal{C}(S)$, is the graph whose set of vertices are the set of isotopy classes of essential simple closed curves, and two vertices are connected by an edge of length 1 if corresponding simple closed curves can be represented disjointly. For $x, y, z \in \mathcal{C}(S)$, the *Gromov product* of y and z with respect to x , denoted $(y \cdot z)_x$ is defined by

$$(y \cdot z)_x := \frac{1}{2}(d(x, y) + d(x, z) - d(y, z)).$$

Since $\mathcal{C}(S)$ is Gromov hyperbolic by [MM], it has the Gromov boundary $\partial\mathcal{C}(S)$. Let $\bar{\mathcal{C}}(S) := \mathcal{C}(S) \cup \partial\mathcal{C}(S)$. A sequence of points $\{x_i \in \bar{\mathcal{C}}(S)\}$ converges to a point $\lambda \in \partial\mathcal{C}(S)$ if $(x_i \cdot \lambda)_x \rightarrow \infty$. We define a *shadow set* by

$$S_x(y, R) := \{z \in \bar{\mathcal{C}}(S) \mid (y \cdot z)_x \geq R\}.$$

Proof of Proposition 4.3. By [Kla, Theorem 1.2 and 1.4] and the fact that $F_+(\omega)$ is uniquely ergodic, we see that $S_x(F_+(\omega), R) \subset U$ as subsets of $\mathcal{PMF}(S)$ for sufficiently large R .

Since $\omega_n x$ converges to $F_+(\omega)$ in $\bar{\mathcal{C}}(S)$, we see that for any D , there exists $N \in \mathbb{Z}_+$ such that for any $n > N$, $\omega_n x \in S_x(F_+(\omega), R + D)$. Hence by the work of Maher [Mah, Proposition 2.13 (5)], for any $\epsilon > 0$, there is ω_n such that $\nu_{\omega_n}(S_x(F_+(\omega), R)) \geq 1 - \epsilon$ where $\nu_{\omega_n}(A) := \nu(\omega_n^{-1}A)$. Note that in [Mah], the measure μ is assumed to have a finite support, however, the finiteness is not used for the proof of results we need here. Hence

$$\nu(S_x(F_+(\omega), R)) \geq \mu_n(\omega_n)\nu_{\omega_n}(S_x(F_+(\omega), R)) \geq \mu_n(\omega_n)(1 - \epsilon) > 0,$$

where μ_n is the n -fold convolution of μ . Since $U \supset S_x(F_+(\omega), R)$, we have $\nu(U) > 0$. \square

4.3. Refinement by cylinders. We finally give neighborhoods U_+ and U_- of $F_+(\omega)$ and $F_-(\omega)$ respectively. We would like to find U_{\pm} so that the vertical lengths of birectangles in $\{\mathcal{R}_n\}$ are bounded from above. Recall that given two $G_{\pm} \in \mathcal{UE}(S) \subset \mathcal{PMF}(S)$, we can construct a birectangle decomposition $\mathcal{R}(G_+, G_-, \mathcal{X}_G)$, where $f : S \rightarrow \mathcal{X}_G$ is a representative of a closest point projection X_G on $\Gamma(G_+, G_-)$ of X_0 . Since the vertical length is independent of representatives of points of the Teichmüller space, we denote birectangle partitions by $\mathcal{R}(G_+, G_-, X_G)$. We abuse notations similarly for τ and η . Let $V(\mathcal{R}(G_+, G_-, X_G))$ denote the maximum of the vertical lengths of birectangles in $\mathcal{R}(G_+, G_-, X_G)$. Since each \mathcal{R}_n is obtained from some $\mathcal{R}(G_+, G_-, X_G)$ by decomposing each birectangle in $\mathcal{R}(G_+, G_-, X_G)$, it suffices to prove the following.

Lemma 4.4. *There exist $V > 0$ and open neighborhoods U_+ and U_- of $F_+(\omega)$ and $F_-(\omega)$ respectively so that the following holds. Suppose that*

(*) $G_{\pm} \in U_{\pm}$ are written as $G_{\pm} = \eta F_{\pm}(\omega)$ respectively for some $\eta \in \text{MCG}(S)$.

Then $V(\mathcal{R}(G_+, G_-, X_G)) < V$.

Proof. Suppose the contrary. Then there are V_n with $V_n \rightarrow \infty$ and $\eta_n \in \text{MCG}(S)$ with $G_{\pm}^n := \eta_n F_{\pm}(\omega) \rightarrow F_{\pm}(\omega)$ such that $V(\mathcal{R}(G_+^n, G_-^n, X_{G^n})) = V_n$. Let $R_n \in \mathcal{R}(G_+^n, G_-^n, X_{G^n})$ be the rectangle with vertical length V_n . Note that since we assume that the total area is equal to 1, the horizontal length of R_n converges to 0. Hence by taking a subsequence if necessary, the vertical boundary $\partial_v R_n$ converges in the Hausdorff topology to an infinite subarc of a singular leaf of $F_-(\omega)$, which

intersects τ_0 only twice. However any infinite singular leaf of $F_-(\omega)$ is dense, so such a singular leaf never exists. \square

Note that since U_+ and U_- are open, we see that $\nu(U_+) > 0$ and $\check{\nu}(U_-) > 0$ by Proposition 4.3. By taking smaller open neighborhoods if necessary, we may suppose that U_{\pm} in Lemma 4.4 satisfy Condition 3.5. From now on, we consider the semi-Markov partition $\{\mathcal{R}_n\}$ constructed with such U_{\pm} and corresponding representation \mathbf{w} of ω whose construction is given in §4.1. We now consider images of cylinders in $\Sigma_A(\mathbf{w}, n)$ by $p(\mathbf{w}, n)$. For notational simplicity, we call the image of cylinders by $p(\mathbf{w}, n)$ cylinders and omit to write $p(\mathbf{w}, n)$. Let $c(n)$ be the number so that the set of cylinders $\mathcal{C}_n(-c(n), c(n))$ refines \mathcal{A} . The existence of $c(n)$ follows from Lemma 3.9. We now prove that $c(n)$ grows sublinearly.

Lemma 4.5. *For the above $c(n)$, we have*

$$\lim_{n \rightarrow \infty} \frac{c(n)}{n} = 0.$$

Proof. It suffices to find $c(n)$ so that the horizontal and the vertical lengths with respect to q'_n of $(-c(n), c(n))_n$ -cylinders are less than $\delta(q'_n)/\sqrt{2}$. Let $V > 0$ be the upper bound given by Lemma 4.4. Then the horizontal and vertical length of any $(-c(n), c(n))_n$ -cylinder is bounded from above by $1/\exp(d_{\mathcal{T}}(w_n X_n, w_{n-c(n)} X_{n-c(n)}))$ and $V/\exp(d_{\mathcal{T}}(w_n X_n, w_{n+c(n)} X_{n+c(n)}))$ respectively. Since the bound for the horizontal length is given similarly, we only discuss the vertical lengths. Let $m := n + c(n)$ and let $\epsilon > 0$ be arbitrary. By Lemma 3.6 and Theorem 3.7, for sufficiently large n , we have $d_{\mathcal{T}}(w_n X_n, w_m X_m) \geq (m - n)L - (n + m)\epsilon$. We also have $\delta(q'_n) \geq 1/\exp(n\epsilon)$ by Lemma 4.2. Let $c^\epsilon(n)$ be the smallest integer with $c^\epsilon(n) > (\log(\sqrt{2}V) + 3n\epsilon)/(L - \epsilon)$. Then for large enough n , $\mathcal{C}_n(-c^\epsilon(n), c^\epsilon(n))$ refines \mathcal{A} . Hence for large enough n , $c(n) \leq c^\epsilon(n)$. Since $\epsilon > 0$ is arbitrary, we have $\lim_{n \rightarrow \infty} c(n)/n = 0$. \square

We are now ready to prove the main theorem.

Proof of Theorem 1.1. Since $\lim_{n \rightarrow \infty} c(n)/n = 0$, there exists $K > 0$ such that $-K < -c(n) + n$ for all $n \in \mathbb{N}$. Therefore we have,

$$\begin{aligned} & \frac{1}{n+1} \log N(\mathcal{A} \vee w_1 \mathcal{A} \vee \cdots \vee w_n \mathcal{A}) \\ & \leq \frac{1}{n+1} \log N(\mathcal{C}_0(-c(0), c(0)) \vee \cdots \vee w_n \mathcal{C}_n(-c(n), c(n))) \quad (\text{by definition of } c(n)) \\ & = \frac{1}{n+1} \log N(\mathcal{C}_0(-c(0), c(0)) \vee \cdots \vee \sigma^{-n} \mathcal{C}_0(-c(n) + n, n + c(n))) \quad (\text{Theorem 1.2}) \\ & \leq \frac{1}{n+1} \log N(\mathcal{C}_0(-K, n + c(n))) \\ & = \frac{\log N(\mathcal{C}(-K, n + c(n)))}{n + c(n) + K} \cdot \left(\frac{n + c(n) + K}{n + 1} \right). \end{aligned}$$

By Lemma 3.11, we have

$$\limsup_{n \rightarrow \infty} \frac{\log N(\mathcal{C}(-K, n + c(n)))}{n + c(n) + K} \leq L.$$

Also by Lemma 4.5,

$$\lim_{n \rightarrow \infty} \left(\frac{n + c(n) + K}{n + 1} \right) = 1.$$

Therefore,

$$\limsup_{n \rightarrow \infty} \frac{\log N(\mathcal{C}(-K, n + c(n)))}{n + c(n) + K} \cdot \left(\frac{n + c(n) + K}{n + 1} \right) \leq L.$$

By putting all estimates together, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log N(\mathcal{A} \vee w_1 \mathcal{A} \vee \cdots \vee w_{n-1} \mathcal{A}) \leq L.$$

Hence the representative \mathbf{w} satisfy that for arbitrary open covering \mathcal{A} , $h(\mathbf{w}, \mathcal{A}) \leq L$. Putting together with Theorem 2.7 and Lemma 2.8, we have $h(\omega) = L$. \square

ACKNOWLEDGEMENT

This work was partially supported by JSPS Research Fellowship for Young Scientists.

REFERENCES

- [AKM] R. L. Adler, A. G. Konheim, M. H. McAndrew. *Topological entropy*. Transactions of the American Mathematical Society 114 (2): 309-319.
- [Ber] L. Bers, *An extremal problem for quasiconformal mappings and a theorem by Thurston*, Acta. Math., 141 (1978), 73-98.
- [CR] Y. Choi and K. Rafi. *Comparison between Teichmüller and Lipschitz metrics*. J. Lond. Math. Soc. (2), 76(3):739-756, 2007. math.GT/0510136.
- [DH] F. Dahmani and C. Horbez, *Spectral theorems for random walks on mapping class groups and $\text{Out}(F_N)$* , arXiv:1506.06790.
- [FLP] A. Fathi, F. Laudenbach, V. Poénaru et al., *Travaux de Thurston sur les surfaces*, Astérisque 66-67, Société Mathématique de France, 1979.
- [FM] B. Farb and D. Margalit, *A Primer on Mapping Class Groups (PMS-49)*, Princeton University Press, 2011.
- [GM] F. Gardiner and H. Masur, *Extremal length geometry of Teichmüller space*, Complex Variables and Elliptic Equations 16.2-3 (1991): 209-237.
- [GK] VK. Gundlach and Y. Kifer, *Random hyperbolic systems* (English summary) Stochastic dynamics (Bremen, 1997), 17-145, Springer, New York, 1999.
- [HM] J. Hubbard and H. Masur, *Quadratic differentials and foliations*, Acta. Math. 142 (1979), 221-274.
- [KM] V. Kaimanovich and H. Masur, *The Poisson boundary of the mapping class group*, Invent. Math. 125 (1996) 221-264.
- [Kar] A. Karlsson, *Two extensions of Thurston's spectral theorem for surface diffeomorphisms*, Bull. Lond. Math. Soc. 46 (2014), no. 2, 217-226.
- [Kla] E. Klarreich, *The boundary at infinity of the curve complex and the relative Teichmüller space*, unpublished manuscript, Ann Arbor 1999.
- [Mah] J. Maher, *Exponential decay in the mapping class group*, J. Lond. Math. Soc. (2) 86 (2012), no. 2, 366-386.
- [MM] H. Masur and Y. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math., 138(1):103-149, 1999.
- [Thu] W. P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bulletin of the American mathematical society 19, (1988) 417-431.
- [Tio] G. Tiozzo, *Sublinear deviation between geodesics and sample paths*, Duke Math. J. 164 (2015), no. 3, 511-539.

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES, THE UNIVERSITY OF TOKYO, 3-8-1 KOMABA MEGURO-KU TOKYO 153-8914, JAPAN

E-mail address: masai at ms.u-tokyo.ac.jp